RISK: Health, Safety & Environment (1990-2002)

Volume 6 | Number 3 Article 4

June 1995

Ranking Risks

Baruch Fischhoff

Ranking Risks*

Baruch Fischhoff**

A Challenging Ideal

Risks are everywhere. How do we decide which ones are worthy of our attention? In an ideal world, on some regular basis, we would review our priorities systematically. That would begin by listing all the risks we face, ordered according to the threat posed by each. It would continue by listing each option for controlling each risk, characterized by some estimate of its effectiveness and cost. It would conclude by identifying the "best buys" in risk reduction, the strategies that achieve the greatest reductions at the least cost. Those costs might be measured in dollars, time, effort, nagging or whatever other resources we have to invest in risk management. As a by-product, this analytical process would leave a list of residual risks, which we cannot reduce at any reasonable price, but which continue to be matters of concern.

In reality, though, such systematic reviews of risk are as rare as systematic reviews of how we spend our time, money or emotions. One obvious constraint on any of these activities is lack of time to perform them. However, even with all the time in the world, there would still be daunting obstacles. Risks are so diverse that it is hard to compile either the list of threats or the set of possible control strategies. We seldom have ready access to credible estimates of the sizes of the risks, the chances for control or the costs of amelioration. Often, the ranking scientific experts know little more.

If we had those figures, we would then have to face difficult tradeoffs. Many of these involve wrenching choices between "your money or your life." More precisely, they ask about our willingness to sacrifice concrete dollars in return for changes in the probability of injury or death. Even if

* Preparation of this manuscript was supported by the National Science Foundation grant SBR-9309428. It benefited from the comments of Dalton Paxman, two anonymous reviewers, and members of the Carnegie Mellon Risk-Ranking Group — Granger Morgan, Lester Lave, Paul Fischbeck, Karen Jenni, Ned Welch, Stephanie Byram, Garrick Louis, Sandra McBride, Laura Painton and Stuart Siegel.

** Dr. Fischhoff is Professor of Social and Decision Sciences and of Engineering and Public Policy at Carnegie Mellon University. He received his B.S. (Mathematics) from Wayne State University and his M.A. and Ph.D. (Psychology) from Hebrew University, Jerusalem.

we knew how to make the tradeoffs posed by individual risk-control options, we would still have to prioritize the different risks competing for the same limited resources. For example, do we invest our research dollars in attempts to reduce the risks of injury, heart disease or cancer? Do we invest our nagging budget in warning our teens about sex, drugs, beer, mixed drinks, AIDS or driving (not to mention skipping school, cheating and getting assignments in late)?

These are not only intellectual challenges. They also confront us with difficult ethical, social and emotional choices. What we do about risks defines us as individuals and as citizens, showing what we value and what we accept as our personal responsibility. Any risk that we neglect can come back to haunt us. Any risk that we face (or explicitly ignore) raises uncomfortable concerns. Any risk that we place on our own agenda may be taken off the plates of others, including those who create and benefit from it.

The next section of this essay offers an account of how individuals might respond to these challenges, whether considering risks under their personal control or ones that they hope to influence through social and political processes. The following two sections consider the role of government in helping citizens to manage risks, first in principle and then in the practical context of recent risk-ranking exercises. The essay then offers a general procedure for risk ranking. It concludes by discussing what can be done with a list of risks.

A Descriptive Account

Faced with such complex problems, we usually just muddle along. We have in place a set of practices that have evolved over time. Some, we have adopted deliberately (e.g., low-salt diets). Others were imposed upon us (e.g., automatic seat belts). Still others were copied from friends with little attention to safety (e.g., smoking) or are of uncertain origin (e.g., triple checking the stove before leaving the house).

Every once in a while, something happens that calls our habits into question. It may make us wonder whether we are needlessly investing in risk control or recklessly leaving ourselves exposed. In many lives, these occasions may come with unnerving frequency. Every Thursday's *New England Journal of Medicine* brings revisions in estimates of some risk's size or some control strategy's effectiveness. Many Tuesdays' *Science Times* brings coverage of more slowly breaking revisions. Most weeks, some TV

news magazine features a health risk or quack cure. Almost every night, local news broadcasts present threats to personal safety, from crime, fire or traffic. Over time, these reports filter into everyday conversation, reaching those who are not news junkies or direct observers. News about risks also arises sporadically in our personal lives, for example, through injuries to friends, reports of asbestos in schools or plans to site nuisances in our neighborhoods (ranging from hazardous waste facilities to half-way houses).

On the positive side, these confrontations offer chances to rethink our priorities. They may force us to think in some depth about unpleasant topics that we might otherwise ignore. They may facilitate collective action or changes in long-standing behaviors. Over time, reviewing the treatment of individual risks should bring our overall priorities in line.

However, having our agenda set in this way has its limitations. One is that nomination may have little to do with the magnitude of risks involved, the usefulness of any new information or opportunities to act. Unfortunately, it is often hard to tell whether a featured risk is worth worrying about. It is rare to find a concise summary of a risk's magnitude or the quality of the underlying science.[1] As a result, citizens must divine the size and certainty of risks from indirect cues. One common, reasonable, but imperfect inference is that if seemingly responsible people raise an issue, then it must be important and they must know something about it. A complementary assumption is that important risks will get reported expeditiously. However, scientists can seize center stage with studies that are important to them personally but that add little to overall understanding; news media often retell familiar stories, while neglecting more serious risks; issues may be ignored just because the story is hard to tell.

Once an issue attracts attention, group processes can take on lives of their own, generating further cues as to the magnitude of a risk. For example, if institutions are perceived as responding callously, citizens may conclude, "if they're so high-handed, they must be hiding something." A forceful public may be construed as strident or hysterial, leading its concerns to be discounted. Risk debates may really be about not losing to the villains (jerks, etc.) on the other side.[2] Whether inadvertent or deliberate,

---

1 Silvo O. Funtowicz & Jerome R. Ravetz, **Uncertainty and Quality in Science for Policy** (1990).er Morgan & M. Henrion, **Uncertainty** (1991); M. Granger Morgan & Max Henrion, **Uncertainty** (1991).

2 **Readings in Risk** (Theodore S. Glickman & Michael Gough, eds. 1990); Sheldon Krimsky & Alonzo Plough, **Environmental Hazards: Communicating Risks as a Social Process** (1988); National Research Council, **Improving Risk Communication** (1989);

poor communication can aggravate whatever "natural" misunderstandings people have about risks.[3]

Moreover, even the best reporting leaves open the question of how important each risk is and how much its control is worth.[4] Better communication will sometimes show dominating alternatives and clear-cut "best buys" in risk reduction. At other times, though, it will just paint a starker picture of a difficult reality. Coping with that reality requires not only processing a lot of information, but also seeing it from many perspectives. It means avoiding the risk of "framing," whereby, in circumstances with limited opportunities for reflection, people's preferences prove sensitive to formally irrelevant aspects of how tradeoffs are described.[5]

Possible Government Roles

Government has some natural advantages in facing these tasks. It has far greater resources than individuals for assembling evidence and analyzing it from diverse perspectives. However, government faces the same obstacles as individuals. It, too, must compile a comprehensive list of risks, including those that are, for whatever reason, commonly ignored. It must summarize the scientific evidence, with adequate representation of uncertainties. For the sake of comparisons, it must render those risks in some common units.[6]

Yet, even with unlimited budgets, government analysts could not solve the risk-ranking problem unambiguously. Each step of the process involves value judgments. It is a question of ethics, not science, to determine which risks can even be considered, how "risk" is to be measured, how the different dimensions of risk should be weighted and how uncertainty

Elaine Vaughan, *Individual and Cultural Differences in Adaptation to Environmental Risks,* 48 Am. Psych. 673 (1993).

[3] *See, e.g.,* Baruch Fischhoff, Ann Bostrom & Marilyn J. Quadrel, *Risk Perception and Communication,* 14 Ann. Rev. Pub. Health, 183 (1993); Kenneth R. Laughery, *Everybody Knows — or Do They?* Ergonomics in Design, Jul. 1993, at 8; Howard Leventhal & Linda Cameron, *Behavioral Theories and the Problem of Compliance,* 10 Patient Ed. & Counsel. 117 (1987); James Reason, Human Error (1990).

[4] Baruch Fischhoff, *Acceptable Risk: A Conceptual Proposal,* 5 Risk 1 (1994); William W. Lowrance, Of Acceptable Risk: Science and the Determination of Safety (1976).

[5] Max H. Bazerman & M. A. Neale, Negotiating Rationally (1992); Robyn M Dawes, Rational Choice in an Uncertain World (1988); Daniel Kahneman & AaronTversky, Choices, Values, and Frames, 39 Am. Psych. 341(1984); Richard H. Thaler, Quasi-rational Economics (1991).

[6] Donald T. Hornstein, *Reclaiming Environmental Law: A Normative Critique of Comparative Risk Analysis,* 92 Colum. L.Rev. 562 (1992).

should be treated.[7] The rankings of risks, and of risk-reduction strategies, will depend on how these issues are resolved. As a result, there will be differences in whose welfare is protected and whose risk-producing activities are restrained. The open nature of government analyses should make such choices relatively transparent. Unfortunately, government analysts often lack explicit legal mandates to make them. Even where there is statutory guidance, it may lack credibility. For example, it is one thing to require cost-benefit analyses, quite another to make people comfortable with value-of-a-life calculations.

Under these circumstances, government analysts can address value issues in several ways. One general strategy is to make as few assumptions as possible, reporting the results in something approaching raw form, assembling the data without digesting them. The work product might include weakly comparable estimates, expressed in different units and accompanied by frank discussions of the sources, assumptions and limitations. Any integration would be left to consumers of the data.

A second strategy is to integrate the evidence in several different ways, each reflecting an alternative value system. Doing so would not prejudge which values are appropriate (among those that are considered). Rather, the analysts would run the numbers under these different assumptions. Relieved of this computational load, citizens could locate themselves in the "space" created by the alternative value systems.[8]

A third strategy is to elicit values from citizens, then derive the analytically appropriate rankings implied by them. Its success depends on these individuals' ability to express their values in the abstract form required by analytical models. That means grappling with difficult ethical questions, using unfamiliar formats and producing public statements. It requires the analysts to carry the citizens along, so that they will see the rankings as expressing the values that they have provided.[9]

Finally, one could allow citizens' panels to determine the rankings with technical staff at their service, explicating the risk data and perhaps suggesting alternative perspectives. The credibility of a citizens' panel

---

[7] Edmund A.C. Crouch & Richard Wilson, Risk/Benefit Analysis (1981); Baruch Fischhoff, Stephen R. Watson & Chris Hope, *Defining Risk,* 17 Policy Sci. 123 (1984).

[8] Lester B. Lave & Hadi Dowlatabadi, *Climate Change Policy: The Effects of Personal Beliefs and Scientific Uncertainty,* 27 Env. Sci. & Tech. 1962 (1993).

[9] Karen E. Jenni, Miley W. Merkhofer & Carol Williams, *The Rise and Fall of a Risk-based Priority System: Lessons from DOE's Environmental Restoration Priority System,* Risk Anal. (in press).

would depend on the populations whose values it represents. If a panel could not converge on a single ranking, it could at least show the array (or disarray) of lay opinion.

Experiments in Ranking

The choice of approach depends, obviously, on the particulars of the situation. For example, Office of Technology Assessment (OTA)[10] has adopted the first approach in its current study of risks to students in school. It has performed the arduous chore of identifying and assembling data from diverse sources. However, it lacks the resources for protracted interactions with citizens (other than Congressional representatives and staff) and the authority to determine a definitive set of values by fiat.

In contrast, the Environmental Protection Agency (EPA)[11] has conducted several internal ranking exercises, as a way of articulating its own priorities. However, participants were restricted to its own staff and advisory committees. Citizens were consulted only through the indirect and imprecise medium of public opinion surveys. EPA concluded that these surveys showed marked discrepancies between the Agency's rankings and those held by the public. If real (and not just methodological artifacts), these discrepancies could reflect differences in perceptions of either scientific or value issues. The sources of such disagreements are, however, notoriously difficult to decode.[12] Surveys typically offer respondents little time to think, opportunity to express complicated thoughts or chance to clarify ambiguous questions.

EPA's response has attempted to bridge this gap through direct interaction with citizens.[13] Specifically, it has promoted state and local risk-ranking exercises. In them, diverse panels of citizens develop rankings of risk, over a period of time and with considerable staff support. These protracted interactions allow participants to mull the complicated issues, clarify apparent differences, recruit needed information and negotiate compromises (if they are to be found). Several dozen such exercises are in

[10] Office of Technology Assessment, Risks to Children in Schools (1995).

[11] Environmental Protection Agency, Unfinished Business: A Comparative Assessment (1987); Environmental Protection Agency, Reducing Risk: Setting Priorities and Strategies (1990).

[12] Baruch Fischhoff, *Risk: A Guide to Controversy,* Appendix to National Research Council, Improving Risk Communications, 211 (1989).

[13] Environmental Protection Agency, A Guidebook to Comparing Risks and Setting Environmental Priorities (1993); Resources for the Future, Setting National Environmental Priorities (1993).

varying stages of planning and completion, producing consensus documents from surprisingly diverse audiences. Their very existence shows a noteworthy willingness to seek compromise.

One critical ingredient in this success has apparently been that each panel is allowed to structure the work as it wishes. The price for conferring such freedom is that there is no common format for the different ranking exercises, no explicit role for analytical procedures and no systematic way to uncover the values underlying the rankings. As a result, it is difficult to generalize results and explicate their rationale. Non-participants may have difficulty accepting rankings derived from unspecified values. Government analysts may lack the explicit guidance needed to translate the rankings into regulations. There may be no way to integrate the results of different ranking exercises.

One Way to Rank Risks

A successful risk-ranking method faces many simultaneous demands. It must reflect the underlying science faithfully, capture the critical dimensions of that science, present that information comprehensibly, secure the input of citizens' values, reach a stable conclusion and convey it credibly to the broad public.

Recently, we proposed a risk-ranking method which, we hope, represents a reasonable compromise among these demands.[14] It is not perfect, nor completely specified. However, it identifies the critical design issues that any deliberate ranking would face and offers an initial approach to them. That is, it provides a task analysis for those who would rank risks, along with one possible response to it. It hopes to capitalize on the successful social process and thoughtful treatment of risk assessment in EPA's risk ranking, while strengthening its analytical core. It was initially produced in response to a request from the Office of Science and Technology Policy for a method that federal agencies could use to prioritize risks, within programs, across programs and across agencies.[15]

Legislation introduced into the 104th Congress (and its immediate predecessors) calls for a dramatic expansion in the use of risk comparisons (e.g., H.R. 9). Whatever form these proposals eventually take, they will need to address the issues raised here. Our procedure involves six steps:

---

[14] M. Granger Morgan et al., A Procedure for Risk Ranking for Federal Risk Management Agencies (Working paper for OSTP 1994).

[15] Ranking Risks, (Paul Portney, ed. in press, Resources for the Future).

1. Define and Categorize the Risks

The universe of relevant risks first must be identified and then must be sorted into a modest number (≤ 30) of roughly comparable risk categories. Those categories should be simple, well-defined, exhaustive, mutually exclusive and sufficiently homogenous that each can be considered in the same light. A rough screening for order of magnitude is needed to avoid creating a few categories lumping extremely large risks — so that the hard work of ranking is merely postponed (until the time comes to deal with the contents of those large categories).

One natural categorization is according to existing regulatory programs. However, those programs may reflect the chaotic bureaucratic and political processes that prompted the need for systematic prioritization in the first place. Categorization might also reflect the *source* of the risk (e.g., power plants), the *agent* of risk (e.g., ozone) or the *failure mode* (e.g., low-level wind shear). Thus, for example, the FAA might decide to categorize risks by source and failure mode. Its sources might be civilian and commercial aircraft, while failure modes might include icing problems, wind shear, engine fires, loss of communication, etc. If so, then one category of risks to be ranked by FAA might be: wind shear accidents involving commercial aircraft.

2. Identify the Relevant Attributes of Risk

Risk is a complex concept. Expected numbers of deaths and injuries are clearly important, but a variety of other considerations, or attributes, may also matter. To the extent possible, these attributes should be: *comprehensive* (to ensure that nothing important has been left out); *non-redundant* (to avoid double-counting); *preferentially independent* (for simpler evaluation procedures); *measurable* (for explicit and consistent estimates) and *minimal* in number (to reduce complexity). Because some criteria (e.g., comprehensive, minimal in number) can conflict, the process of choosing attributes will have to involve judicious compromises.

The compromise that we proposed characterizes each risk according to the three dimensions of risk that have emerged in psychometric studies of perceived risk:[16] *number of people affected*, *knowledge* and *dread*. Because those studies focused on risks to humans, *ecological impact* was added as a fourth dimension. For any given hazard, people's judgments about attributes

[16] Baruch Fischhoff et al., *How Safe is Safe Enough? A Psychometric Study of Attitudes Towards Technological Risks and Benefits*, 8 Policy Sci. 127 (1978); Paul Slovic, *Perceptions of Risk*, 236 Science 280 (1987); Paul Slovic, Baruch Fischhoff & Sarah Lichtenstein, *Rating the Risks*, 21(4) Environment 14-20, 36-39 (1979).

that fall *within* one of these sets tend to show high inter-attribute correlations. In contrast, comparisons of attributes that lie in different sets display low inter-attribute correlations. Thus, as long as the attributes used in the ranking process include a few attributes from each set, the results of a ranking process should not to depend very much on which specific attributes are used.

We propose that rankers select two markers for each dimension, from the large set of highly correlated possibilities. For example, "dread" could also be measured by individual controllability, catastrophic potential and outcome equity (among other possibilities). This procedure will allow rankers to choose marker attributes that they find meaningful, while ensuring that the rankings produced by groups using different markers would be similar (because of the correlations among alternative markers and the coverage of the four dimensions). Our full proposal offers operationalizations of 4-6 possible markers for each dimension.

Table 1[17]

A Possible Quantitative Summary of Risks

Number of people affected	*Degree of environmental impact*	*Knowledge*	*Dread*
Annual expected number of fatalities 0 – 450 – 600 (10% chance, zero)	Area affected by ecosystem stress or change 50 km²	Degree to which impacts are delayed 1 – 10 years	Catastrophic potential 1000 x expected annual fatalities
Annual expected number of person-years lost 0 – 9000 – 18000 (10% chance, zero)	Magnitude of environmental impact modest (15% chance, large)	Quality of scientific understanding medium	Outcome equity medium (ratio = 6)

3. Describe the Risks

Once the categories and attributes are set, technical staff would summarize the scientific evidence for each. In a standard format that summary would include (a) a qualitative description of the risk, (b) a quantitative evaluation of the risk in terms of each chosen attribute and (c) a brief description of the state of scientific understanding (expressing uncertainty, broadly defined). Table 1 shows a possible format for the

17 Source: Morgan et al., *supra* note 14.

quantitative summary. To accommodate rankers with diverse educational background, the narrative summaries should be no more technical that *Popular Science.* To facilitate making comparisons, the summary forms should be small enough to be easily handled and sorted. For example, one might use legal-sized sheet of paper, turned sideways and folded in half, with the tabular summary of Table 1 appearing on the cover page under a brief description of the risk and followed by the narrative. Technical staff would be available to provide whatever additional detail is needed.

4. Select the Groups of Rankers

The rankers should represent those citizens whose values are to be captured. In our worked example, focused on establishing priorities for federal agencies, we proposed four independent groups: one with federal agency risk personnel, one with state and local risk managers and two groups of laypeople. Each group would reflect the diversity of its underlying population. Their opinions would be interpreted as the conclusions that similar citizens would reach were they to invest similar effort in these topics. Membership would be limited to 10-15, in order to allow for active participation by all. Groups would manage their own affairs, with staff support. They would select their own chair and vice-chair, who would receive suitable training in the procedure and in group process. Multiple groups are used to increase confidence that the results are robust and not just the product of particular group dynamics. Constituting groups with similar levels of technical expertise (about risks and about regulation) is intended to promote interaction among equals.

5. Perform the Rankings

We propose a series of four meetings, during which each group would seek a consensual ranking. Before beginning this process in earnest, members would individually evaluate the risks using a simplified multi-attribute weighting approach. Toward the end of a group's sessions, members would review its tentative conclusions in the light of these initial individual rankings. This form of triangulation is intended to protect individuals against framing effects, by asking them to reconcile two potentially different ways of looking at the problem. It is intended to protect groups against dominating personalities or collective myopia by giving equal standing to each members' initial position. Upon completing its work, each group would select representatives to an intergroup synthesis meeting. This meeting would seek the maximum consensus that is possible, within the constraints that the constituent groups set for their representatives. The

final agreement, ratified by the four groups, would be announced with suitable fanfare.

In all these deliberations, the emphasis is on sorting the risk categories into a few broad classes, paying particular attention to identifying those risks that deserve the highest — and the lowest — ranks. The groups should not expend energy on the meaningless task of precisely ordering risk categories whose ranks broadly overlap. The greatest benefit of the whole exercise is likely to come from identifying risks with clear ranks, especially ones that are not commensurate with the resources invested in their management.

6. Provide a Reasonably Rich Description

The same summary ranking can mean quite different things if it reflects a strong consensus or a weak plurality of views. It can motivate different actions if residual disagreements reflect conflicting values or alternative interpretations of uncertain scientific evidence. As a result, an appropriate summary is needed, capturing these sources of disagreement, as well as any problematic procedural issues. While a clear consensus may be needed to break political deadlocks, clearly characterized disagreements can still focus future research and debate.

Our full proposal considers various other issues, such as how to balance the confidentiality needed for frank discussions with the openness needed for credibility with nonparticipants (as well as complying with open-meeting laws). It also identifies unresolved issues and highlights the need to evaluate procedures before implementation.

What Can One Do with a Risks Ranking?

In a sense, risk ranking can do no more than satisfy curiosity. A list of risks carries no necessary implications for action. Big risks might be neglected if nothing could be done; small risks might be reduced if that could be done cheaply. Ultimately, one wants to rank not risks, but actions, to identify the best buys in risk reduction.

Public risk-ranking processes have typically stopped short of recommending actions. In part, this has occurred because those conducting them lacked the authority to go further. For example, OTA was not asked to determine what to do about risks to students in school; EPA's voluntary consensus-building might have collapsed had they attempted to take the next step.

In principle, this is a disappointing conclusion to such ambitious efforts. In practice, though, consensual risk rankings can make a difference — even without solving the political problems of securing a mandate for change or the intellectual problems of mapping risks to actions. If a risk is clearly small, cheap controls should not be foregone. Yet, neither should they be sought too actively. If a risk is clearly large, it deserves attention unless there is some immutable reason why it cannot be reduced. At times, there will be a simple risk-action connection, so that consensus on the former will carry to the latter. If several small risks obviously have common treatment, they could be grouped (and moved up the list).

At times, it may become apparent that there is no way to transfer resources from overmanaged risks to undertreated ones. Such situations can provoke anger insofar as some implicit promise of fungibility underlies the exercise. (Why compare risks if nothing can be done about reordering their priorities?) Meaningless risk comparisons are widely held to be a source of public anger at risk managers.[18] Yet, such anger may be needed to create a mandate for change. An agency frustrated with its enabling legislation might even take the calculated gamble of inviting public anger, hoping that the resulting furor will position it better for the long run.

The dangers with provoking anger at mistaken priorities is that the resulting turmoil might throw out the good with the bad. Times of change are often times for mischief, with those closest to the seats of power attempting to settle private accounts under the banner of public reform. Our proposal is intended to promote the possibility of orderly change, by increasing confidence that it can be done in an open, regulated and scientifically credible way. Its underlying article of faith is that a well-managed, mutually respectful process will reveal some significant areas of agreement, even among diverse individuals. The result will be fewer, but better focused, conflicts than would arise without such an opportunity.

[18] Vincent T. Covello, Paul M. Sandman & Paul Slovic, Risk Communication, Risk Statistics, and Risk Comparisons: A Manual for Plant Managers (Chem. Mfrs Assn. 1988); Baruch Fischhoff et al., Acceptable Risk (1981).